



Published in final edited form as:

Acad Emerg Med. 2015 November ; 22(11): 1298–1306. doi:10.1111/acem.12814.

Comparison of Two Sepsis Recognition Methods in a Pediatric Emergency Department

Fran Balamuth, MD, PhD, MSCE, Elizabeth R. Alpern, MD, MSCE, Robert W. Grundmeier, MD, Marianne Chilutti, MS, Scott L. Weiss, MD, Julie C. Fitzgerald, MD, PhD, Katie Hayes, Warren Bilker, PhD, and Ebbing Lautenbach, MD, MPH, MSCE

Department of Pediatrics (FB, RWG), Department of Anesthesia and Critical Care (SLW, JCF), Center for Clinical Epidemiology and Biostatistics (WB, EL), Department of Medicine, Division of Infectious Diseases (EL), University of Pennsylvania Perelman School of Medicine, Philadelphia, PA; Division of Emergency Medicine and Center for Pediatric Clinical Effectiveness (FB, KH), Center for Biomedical Informatics (RWG, MC), Division of Critical Care (SLW, JCF), Children's Hospital of Philadelphia, Philadelphia, PA; Department of Pediatrics, Northwestern University Feinberg School of Medicine and Division of Emergency Medicine, Ann and Robert H. Lurie Children's Hospital of Chicago, (ERA) Chicago, IL

Abstract

Objectives—To compare the effectiveness of physician judgment and an electronic algorithmic alert to identify pediatric patients with severe sepsis/septic shock in a pediatric emergency department (ED).

Methods—This was an observational cohort study of patients older than 56 days with fever or hypothermia. All patients were evaluated for potential sepsis in real time by the ED clinical team. An electronic algorithmic alert was retrospectively applied to identify patients with potential sepsis independent of physician judgment. The primary outcome was the proportion of patients correctly identified with severe sepsis/septic shock defined by consensus criteria. Test characteristics were determined and receiver operating characteristic (ROC) curves were compared.

Results—Of 19,524 eligible patient visits, 88 patients developed consensus-confirmed severe sepsis or septic shock. Physician judgment identified 159, and the algorithmic alert identified 3,301 patients with potential sepsis. Physician judgment had sensitivity of 72.7% (95% CI = 72.1% to 73.4%) and specificity 99.5% (95% CI = 99.4% to 99.6%); the algorithmic alert had sensitivity 92.1% (95% CI = 91.7% to 92.4%), and specificity 83.4% (95% CI = 82.9% to 83.9%) for severe sepsis/septic shock. There was no significant difference in the area under the ROC curve for physician judgment (0.86, 95% CI = 0.81 to 0.91) or the algorithm (0.88, 95% CI = 0.85

Address Correspondence to: Fran Balamuth MD, PhD, MSCE, Division of Emergency Medicine, Children's Hospital of Philadelphia, 3501 Civic Center Blvd, CTRB 9206, Philadelphia, PA 19014; Phone: 215 590 7295, balamuthf@email.chop.edu.

Presentations: Pediatric Academic Societies Annual Meeting, Vancouver, BC, May 2014.

Disclosures: FB received career development support from NIH NHLBI K12-HL109009. SW received support from NIH NICHD K12-HD047349 EL received support from the Centers for Disease Control and Prevention Epicenters Program grant U54-CK000163. Funders were not involved in design and conduct of the study; collection, management, analysis, interpretation of the data; preparation, review, or approval of the manuscript. The authors have no other financial relationships relevant to this article to disclose. Dr. Alpern, an associate editor at this journal, had no role in the peer-review process or publication decision for this paper.

to 0.91; $p = 0.54$). A combination method using either positive physician judgment or an algorithmic alert improved sensitivity to 96.6% and specificity to 83.3%. A sequential approach, in which positive identification by the algorithmic alert was then confirmed by physician judgment, achieved 68.2% sensitivity and 99.6% specificity. Positive and negative predictive values for physician judgment vs. algorithmic alert were 40.3% vs. 2.5% and 99.88 % vs. 99.96%, respectively.

Conclusions—The electronic algorithmic alert was more sensitive but less specific than physician judgment for recognition of pediatric severe sepsis and septic shock. These findings can help to guide institutions in selecting pediatric sepsis recognition methods based on institutional needs and priorities.

INTRODUCTION

Sepsis is a complex clinical syndrome resulting from a systemic inflammatory response to infection. Each year, there are approximately 72,000 children hospitalized for severe sepsis in the United States, resulting in significant morbidity and mortality, and nearly \$4.8 billion in U.S. health care expenditures.^{1–3} There have been significant advances in early recognition and overall approach to sepsis over the past decade, which have demonstrated improved patient outcomes with protocol-driven treatment in patients with sepsis.^{1–5}

Pediatric studies performed in the intensive care setting have demonstrated an association of delayed or inadequate goal-directed resuscitation with increased mortality.^{2,3} In response to these findings, the implementation of protocol-driven management for pediatric patients with sepsis in the emergency department (ED) has been able to expedite care, reduce hospital length of stay (LOS), and decrease mortality.^{4–6}

An important limitation to widespread implementation of protocol bundles for pediatric sepsis is the challenges of early and accurate identification of patients with potential sepsis who may benefit from these intensive therapies.⁷ This identification process is particularly problematic in the pediatric ED, where there is a high prevalence of fever with signs of the systemic inflammatory response syndrome (SIRS), despite the fact that the vast majority of these patients are not seriously ill.^{8,9} Moreover, hypotension tends to be a late finding in young children, and is less useful to drive early recognition and treatment in pediatric sepsis.⁸

The electronic health record (EHR) is a powerful platform on which decision support tools can be developed and implemented to expedite the sepsis recognition process. The EHR incorporates vital signs, physical exam findings, and past medical history that can be filtered through an electronic algorithm to alert clinicians for suspicion of sepsis.¹⁰ One such sepsis alert using a combination of SIRS criteria and hypotension reduced time to key resuscitative interventions for adults with severe sepsis and septic shock in a general ED setting.¹¹

We developed an electronic algorithmic alert based on key elements outlined by the American Academy of Pediatrics (AAP) Septic Shock Collaborative¹² to identify patients with potential sepsis in a large, busy academic pediatric ED. Elements included in the alert are standard (age-based vital signs, perfusion, mental status, underlying high-risk

conditions) and should be generalizable outside of a dedicated pediatric setting. In this study, we sought to determine the performance characteristics based on a reference standard of an electronic algorithmic alert, routine physician judgment, and both a combination and a sequential mechanism using the algorithmic alert and physician judgment.

METHODS

Study Design

This was a retrospective cohort study. The Institutional Review Board at The Children's Hospital of Philadelphia (CHOP) approved the protocol.

Study Setting and Population

This study was performed at a single quaternary care urban children's hospital with approximately 90,000 ED visits and 3,600 pediatric intensive care unit (PICU) admissions annually.

All visits for patients between 56 days and 18 years old to the CHOP ED between January 1, 2012 and May 31, 2013 with documented fever $\geq 38.5^{\circ}\text{C}$, documented hypothermia $< 36.0^{\circ}\text{C}$, or chief complaint of fever were included. Eligible visits were identified through the CHOP fever registry, an extensive database that captures all ED patient visits with documented fever $\geq 38.5^{\circ}\text{C}$, documented hypothermia $< 36.0^{\circ}\text{C}$, or chief complaint of fever. Because a putative prospective sepsis screening tool would be used at the visit level, we used ED visit as the unit of analysis.

Study Protocol

The CHOP fever registry contains over 150 patient data elements for each visit extracted from the EHR, including demographic data, past medical history, ED vital signs, ED nursing assessment of capillary refill, pulses, skin condition, mental status, ED and hospital laboratory results, ED and hospital therapies, and ED and hospital LOS.⁶

Briefly, severe sepsis is defined as SIRS plus potential infection plus at least two organ dysfunctions, and septic shock is defined as SIRS plus potential infection plus cardiovascular system dysfunction. Organ dysfunction definitions are listed in Data Supplemental 1. This was the only variable that was not automatically extracted from the EHR and was determined by medical record review. For ED patients admitted to the PICU within 24 hours of ED triage, the outcome was determined using a standardized daily screening checklist of all PICU patients that was completed by the PICU team as part of routine clinical care. For ED patients who either died prior to PICU admission or for whom organ dysfunction completely resolved prior to transfer from the ED, we reviewed the medical record to identify patients who met consensus criteria for severe sepsis/septic shock in the ED but who no longer met criteria for severe sepsis/septic shock after hospital admission. Prior to review, all investigators concurred on the application of published definitions of severe sepsis/septic shock using medical record review. Final determination of severe sepsis/septic shock was determined using all clinical and laboratory data available in the medical record by three investigators (FB, JF, SW) blinded to the algorithm alert but not

to physician judgment and was documented on a standardized report form within the study database. Any discrepancies in outcome determination were resolved by consensus among the three investigators during monthly meetings. This occurred as part of an ongoing quality improvement project and investigators at that time were blinded to the hypothesis in this study.

Two strategies of sepsis recognition were studied: routine physician judgment (physician judgment) and an electronic algorithmic approach (the algorithm alert). Physician judgment was determined to be positive if the treating clinical team initiated treatment using the existing ED sepsis pathway. Treatment on the ED sepsis pathway requires an electronic physician order that is a reliable and objective indicator available for all ED visits at our institution. Physician judgment occurred prospectively as part of routine clinical care. Before and throughout the study period, educational sessions were provided to instruct physicians on how to recognize severe sepsis/septic shock and when to initiate the ED sepsis pathway. Physicians were directed to take the following items into consideration when considering initiating the sepsis pathway: vital signs, past medical history, mental status, and perfusion. The educational sessions did not instruct physicians to assign particular value to any sepsis risk factors, but instead instructed them to synthesize their clinical opinion in determining their final decision.

The algorithm alert incorporated demographic, clinical, and physical parameters recommended by the American Academy of Pediatrics Septic Shock Collaborative (Data Supplement 2).¹² We retrospectively applied the algorithm alert to the eligible study cohort within the CHOP fever registry. The algorithm alert was first applied following completion of the nursing triage assessment with subsequent reconsideration at any point in which new vital signs or nursing assessments were entered into the EHR, thus providing a continual screen throughout the patients' ED visit. The algorithm alert was positive for potential sepsis if 1) an abnormality was noted in at least three of the following vital sign categories: temperature, heart rate, respiratory rate, or blood pressure recordings; or 2) abnormalities in any two vital sign categories^{13,14} plus at least one of the following: 1 high-risk condition (Data Supplement 2), abnormal perfusion (defined as capillary refill > 2 seconds or abnormal pulses), or abnormal mental status (defined as inconsolable, eyes do not open to stimuli, lethargic, agitated, or non-responsive). All of the algorithm alert variables were electronically extracted directly from the EHR: vital signs from nursing flowsheets, high-risk conditions from the problem list, and perfusion and mental status assessments from automated prompts within the nursing flowsheet rows routinely completed during the triage process. The EHR is built such that physical assessment fields for perfusion and mental status default to normal and are actively changed by the triage nurse if abnormal. These settings are standard at our institution and were not specifically formatted for this study. While abnormal vital signs could occur at any point during the ED visit, the algorithm alert only considered high-risk conditions, perfusion, and mental status assessments that were available following the initial triage assessment. The alert required vital signs to be from the same time point to trigger a positive screen (i.e. tachycardia and tachypnea needed to be simultaneous; the alert would not take heart rate from one time point and respiratory rate from another time point to yield a positive alert).

In addition to consideration of the algorithm alert and physician judgment alone, we a priori planned to determine the role for two different blended strategies. For the combination method, the alert was positive for potential sepsis if patients had either a positive physician judgment or the algorithm alert and negative if both physician judgment and the algorithm alert were negative. For the sequential method, the alert was positive for sepsis if an initially positive algorithm alert was subsequently confirmed by a positive physician judgment. The combination method simulated a scenario in which either the physician judgment or the algorithm alert would lead to therapy for potential sepsis, whereas the sequential method simulated a more common clinical practice in which the algorithm alert would be confirmed or refuted by physician judgment prior to initiation of therapy for potential sepsis.

Data elements that were collected electronically directly from the EHR onto a standardized reporting form in REDCap included age, sex, race, vital signs, laboratory testing, medications given, patient disposition, length of stay, and complex chronic conditions (CCC). The CCC classification scheme uses a validated grouping of ICD9-CM codes to categorize comorbid disease processes into the following nine categories: malignancy, hematology/immune, respiratory, gastrointestinal, metabolic, neuromuscular, cardiovascular, renal, and other congenital abnormalities.¹⁵

Outcomes

The primary outcome was development of severe sepsis or septic shock within 24 hours of ED triage time as defined by consensus guidelines.¹³

Data Analysis

Statistical analyses were conducted using Stata 12.1. Continuous variables are summarized as the median with interquartile range (IQR) and compared using the Wilcoxon rank sum test. Categorical variables are presented as proportions and compared using the chi-square test. The area under the receiver operating characteristic curve (AUC) for physician judgment, the algorithm alert, combination method, and sequential method to identify visits for patients with severe sepsis/septic shock were compared using the method described by DeLong et al.¹⁶ We calculated the sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios with associated 95% confidence interval (CI) for each identification method. We a priori planned stratified analyses to evaluate the test performance of physician judgment, the algorithm alert, combination method, and sequential method by age and the presence of at least one CCC. To account for possible inaccuracy in perfusion and mental status assessment variables due to the EHR defaulting these variables to “normal,” we conducted sensitivity analyses to assess for the contribution of these variables in alert performance. We compared the proportion of visits for patients with severe sepsis/septic shock with positive algorithm alert due to vital signs alone versus a positive algorithm alert that included abnormal perfusion or mental status assessment. To evaluate the possibility of misclassifying a patient with normal vital signs but who had inaccurate default nursing documentation of normal mental status and perfusion, we reviewed the medical record to determine if there was any physician documentation of abnormal perfusion or mental status in the ED of all cases of confirmed severe sepsis/septic shock that had negative algorithm alerts. Finally, we used multivariable logistic regression

to test the strength of association for each component of the algorithm alert with the reference standard outcome of severe sepsis/septic shock. P-values < 0.05 were considered statistically significant.

RESULTS

During the 17-month study period, there were 138,979 total visits to the ED of which 19,524 (representing 19,124 unique patients) met inclusion criteria for fever, hypothermia, or chief complaint of fever. Study patients were younger (median 2.1 years, IQR 1.1 to 4.6 years) than the overall ED population (median 4.7 years, IQR 1.8 to 10.4 years) during the study period ($p < 0.001$), but were similar in sex and race (both $p > 0.05$). The overall characteristics of study patients and the general ED population are shown in Table 1.

Of the 19,524 eligible patient visits, 159 (0.8%) had positive physician judgment and 3,301 (16.9%) triggered positive algorithm alerts for potential sepsis. When considering the two combined methods, 3,334 (17.1%) triggered the combination method and 126 (0.6%) triggered the sequential method. The characteristics of patient visits comprising the positive physician judgment, the algorithm alert, the combination method, and the sequential method groups are shown in Table 1b. The admission rate to the PICU was 47.8% for physician judgment-positive patients ($p < 0.001$), and 6.5% for algorithm alert-positive patients. Only 2.5% of physician judgment-positive patients, compared to 43% of the algorithm alert-positive patients ($p < 0.001$), were discharged home from the ED.

We catalogued the amount of missing data for each algorithm alert-related variable, which is detailed in Data Supplement 3. There was missing data for heart rate and respiratory rate in < 0.1% of subjects, and there was no missing data for capillary refill or mental status. There were missing blood pressure values for 33% of subjects. We performed several sensitivity analyses to evaluate the effect of this missing data. Of the 6,399 subjects with missing blood pressure values, 4743 (74%) had either no or one abnormal vital sign with normal mental status and capillary refill. For these subjects, even if the blood pressure had been abnormal, they still would have been classified as algorithm alert negative. There was one subject with the severe sepsis/septic shock outcome in this group, who was a complex patient who looked well with normal vital signs in the ED who decompensated on the floor 12 hours later. There were 1,476 subjects with two abnormal vital signs where an abnormal blood pressure value could have changed the result of the algorithm alert. Of these, one patient had the severe sepsis/septic shock outcome, and this patient was the algorithm alert positive due to underlying high-risk condition.

Eighty-eight (0.45%) patients of the overall eligible cohort met consensus-defined criteria for severe sepsis or septic shock within 24 hours of ED triage time. The test characteristics of each identification strategy are detailed in Table 2. The physician judgment method had sensitivity of 72.7% and specificity of 99.5%. The algorithm alert had sensitivity of 92.1% and specificity of 83.4%. When the alerts were used in combination, considering either a positive physician judgment or a positive algorithm alert, sensitivity was 96.6%, and specificity was similar to the algorithm alert alone at 83.3%. The sequential method had a sensitivity of 68.2% and specificity of 99.6%. The combination method strategy achieved

the highest area under the receiver operating characteristic curve, at 0.90. We also determined likelihood ratios and positive and negative predictive values, which are presented in Table 2. Figure 1 shows visit counts within each cell of the 2x2 tables for each test.

We evaluated the primary ED diagnosis and disposition of the 21 patients with severe sepsis/septic shock who were identified by the algorithm alert but not by physician judgment (Table 3a). In addition, we compared process and outcome measures in patients treated according to the ED sepsis protocol/order set to those who were not. We found that sepsis protocol patients had shorter median times to antibiotics, as well as shorter median ICU LOS, and hospital LOS, compared to non-protocol patients (Table 3b).

Table 4 summarizes the age- and CCC-stratified analyses for physician judgment and the algorithm alert. Notably, sensitivity of the algorithm alert was highest in the oldest age strata and was lowest in the 1–4 year old age strata. The physician judgment method had its greatest sensitivity in the youngest age group, and consistently high specificity across all age groups (Table 4a). Table 4b shows differences in physician judgment and the algorithm alert performance in the presence or absence of at least one CCC. The physician judgment method had similar sensitivities and specificities in patients with or without CCCs. The algorithm alert method had increased sensitivity but decreased specificity in patients with at least one CCC compared to patients with no CCC.

We performed sensitivity analyses to account for the fact that mental status and perfusion assessment variables are automatically defaulted to “normal” in the EHR and thus may have been prone to misclassification bias. First, we determined that 3,242 of the 3,301 (98.2%) positive alerts were attributable to vital sign criteria and high-risk conditions alone, even when mental status and perfusion assessments were ignored. The remaining 59 patients (0.8%) had positive algorithm alerts due to either abnormal mental status or perfusion, with two of these 59 patients ultimately meeting criteria for severe sepsis/septic shock. The proportion of patients with positive algorithm alerts who met consensus criteria for severe sepsis/septic shock was not different when considering a positive algorithm alert due to vital signs/high-risk conditions alone or due to a documented abnormality in mental status or perfusion (2.7% vs. 3.3%; $p = 0.8$).

In multivariable logistic regression analysis (Table 5), the components of the algorithm alert that were significantly associated with the reference-standard outcome of severe sepsis/septic shock were hypotension, mental status or perfusion assessment, and abnormal heart rate.

DISCUSSION

We have demonstrated that an electronic algorithmic approach for identifying pediatric patients with potential sepsis based on abnormal age-based vital signs and at least one high-risk condition, abnormal perfusion, or abnormal mental status, has higher sensitivity but lower specificity than physician judgment in a cohort of pediatric ED patients with fever or hypothermia. Algorithm alert performance was affected both by age and presence of

comorbid conditions. We observed the highest sensitivity when a combination of both the physician judgment and the algorithm alert methods was used. The specificity of the physician judgment method highlights the critical importance of bedside assessment in sepsis recognition and underscores the importance of clinician assessment in the initiation of sepsis protocols, and not relying solely on electronic screens. In addition, the PPV of the algorithm alert was quite low (2.5%), likely reflecting the low prevalence of severe sepsis in this cohort. Because severe sepsis in children will likely always be a rare outcome in any given population at risk, screening methods are unlikely to have a robust PPV.

Determining which alert is “best” practice is a complex and difficult decision and may differ for distinct institutions. We identified that the algorithm alert studied here has advantages over physician judgment in terms of sensitivity, which potentially should be prioritized over specificity in sepsis screening and clinical decision-making because of the high cost of missing a single case of sepsis, where timely treatment has been shown to improve outcomes.^{11,17} However, the physician judgment method clearly outperformed the algorithm alert in terms of specificity. The increased specificity afforded by the physician judgment method is also important as we consider the dangers of unnecessary treatment, including antibiotic overuse, resource utilization, and trauma to children experiencing unnecessary intensive interventions. In addition, it is important to be mindful about “alert fatigue” associated with electronic alerts with high signal-to-noise ratios.¹⁸ Based on the data we show here, with the highest observed sensitivity by the combination method, our institution plans to implement a prospective alert utilizing a combination of both tests, where an electronic alert is implemented along with bedside patient assessment utilizing physician judgment, thus taking advantage of each test where most useful. However, another institution may choose to prioritize specificity based on their identified sepsis needs. This “either/or” method would have allowed us to identify four patients with severe sepsis/septic shock who were not identified by the algorithm alert method but were identified by physician judgment, indicating that there is additional valuable information provided by physician judgment that could be missed if only the algorithm alert-positive patients were subsequently evaluated using physician judgment. The sequential method presented in this study, where only the algorithm alert-positive patients are evaluated by physician judgment, displayed lower sensitivity. It is important to note, however, that the sequential method simulated here was not truly sequential due to the retrospective application of the algorithm alert; the treatment team was not aware of the algorithm alert result when making a clinical decision, which could have affected the treatment decision.

The value that the algorithm alert adds in this study is that it was able to identify patients who were not identified by physician judgment. The initial importance of this is underscored by our demonstration that the patients who were “missed” by physician judgment had longer times to antibiotic therapy, as well as longer ICU and hospital LOS. Because of the high financial costs and unmeasurably high personal costs due to these morbidity- and, potential, mortality-related outcomes in children, we consider that most centers will aim to prioritize sensitivity in the case of pediatric sepsis recognition over specificity. That said, it is important to understand the bedside response to a positive algorithm alert in a putative prospective alert because this will help to mitigate the low specificity of the algorithm alert, and thus increase the sensitivity: specificity ratio, or “bang for your buck”, in contemplating

instituting a prospective alert. In real time, the goal of this type of alert would be to bring the medical team to the bedside with the specific goal of evaluating the need for aggressive sepsis therapy. Importantly, the goal of such an alert would not be to treat all alert-positive patients with IV antibiotics and fluids, but more to raise the possibility of a sepsis diagnosis to the treating team to ensure that it is within their differential diagnosis and that they act accordingly based on bedside judgment. Implementation of such a protocol may be more challenging in a center with low pediatric volume and thus less clinician pediatric experience. However, we feel that initial implementation in pediatric centers is an important starting point, and future studies will aim to study these centers and develop specific interventions to increase the generalizability of these screening practices.

As there is increasing pressure both from hospitals as well as legislative bodies in some states to put systems in place to improve sepsis recognition,^{19–21} it is essential to study proposed methods to ensure that they function as envisioned. To our knowledge, one pediatric sepsis alert has been studied to date, which defined a positive alert as temperature-corrected tachycardia, the presence of high-risk conditions, or clinically ill appearance.²² Reported sensitivity of the alert in this study was 81%, with 89% specificity. However, this previous study defined the shock outcome as any patient for whom the clinician decided to activate the shock protocol, not the organ failure-based sepsis definitions we used, and thus could be subject to misclassification bias in the measurement of the primary outcome. In addition, unlike our study, this previous study used only the triage set of vital signs, thus abnormalities that developed later in the ED stay may have been missed. Indeed, in a study by Paul et al., only 43% of children with severe sepsis/septic shock met criteria for this diagnosis immediately on presentation, with 58% progressing to severe sepsis/septic shock later in their ED courses.¹⁷ The generalizability of the algorithm alert in this manuscript is currently being evaluated by the AAP Septic Shock Collaborative, in which different implantation strategies for similar but not identical alerts are being undertaken by 21 hospitals nationwide in a variety of clinical settings.¹²

Several infrastructural advantages at our center allowed us to carry out this work in a largely automated fashion. Our bioinformatics infrastructure enabled us to automatically extract clinical data from the EHR in a rigorous manner. In addition, collaborative sepsis care infrastructure built jointly by ED and PICU teams allowed us a mechanism to identify patients with sepsis who may have been missed by current ED identification methods.

The alterations in algorithmic performance by age underscore the unique challenges of sepsis recognition across the age spectrum. We were surprised to note the low sensitivity of the algorithm alert in the 1 to 4 year old age group, as well as the low specificity in the adolescent age group. One possible explanation for these findings is that the vital sign cutoffs in the algorithm alert were generated a priori by the AAP Septic Shock Collaborative and future work may help determine other or modify cut points using HER-based resources in an evidence-based fashion.

It is also interesting to note that the negative predictive value of both identification methods was very high (>99.8%). Although this is likely most indicative of the low prevalence of sepsis in the population of children with fever, it is notable that children who had neither a

positive algorithm alert nor physician judgment test for potential sepsis using the combination method were very unlikely to have severe sepsis or septic shock (NPV 99.98%).

LIMITATIONS

We retrospectively applied the electronic algorithm alert to our patient cohort, and there may be important performance differences if it were used as a real-time electronic alert in clinical practice. Because of the high prevalence of SIRS criteria in the pediatric population and concerns for a high false positive rate, an initial retrospective evaluation of the algorithm alert was necessary to understand and optimize its performance prior to prospective validation and implementation. The knowledge gained from studying the algorithm's performance prior to prospective implementation is a critical step to optimizing an evidence-based alert. This retrospective application of the algorithm alert likely explains the surprising finding that the sequential method in our study had worse sensitivity than the algorithm alert alone. Had the algorithm alert been implemented in real time and a positive screen truly triggered bedside physician evaluation, test characteristics of the sequential method would have likely improved.

Since the algorithm alert was limited to data elements within the EHR, this strategy may miss patients if vital signs and/or nursing assessments are not recorded in the EHR in a timely and accurate manner. During this study period in our institution, patients who were treated only in the ED resuscitation room did not have real-time electronic flowsheet charting available. Although all of these patients would have triggered the algorithm alert if data were entered into the EHR in this study, this highlights one important pragmatic challenge of relying on an electronic alert. As institutions are actively changing over to real-time EHR documentation in the resuscitation room, this limitation will become less of a concern, but computer downtimes will continue to be problematic. However, an important goal of an electronic alert is to identify the difficult to recognize patient, and the patient who is already in the resuscitation room is not at such great risk of missed identification.

Although the perfusion and mental status findings were abstracted from nursing flowsheets in the EHR, these values are defaulted to normal, and it is possible that there were some patients with abnormal assessments who were not captured in this data set. This did not appear to affect the results based on our sensitivity analyses.

Although physician judgment was made independent of (and prior to) the algorithm alert, it is possible that physician judgment influenced nursing assessment, which could have then affected the retrospectively applied algorithm alert trigger. Also, if the ICU team was more likely to screen a patient as having severe sepsis or septic shock because he or she was treated on the ED protocol, this would bias the study towards an overestimate of physician judgment performance, and thus towards the null hypothesis. Similarly, because physician judgment occurred as part of usual clinical care and use of the sepsis protocol could have been documented in the medical record, we were not able to blind reviewers to physician judgment in terms of outcome determination. However, the assessment of the consensus criteria outcome is very prescribed and based on objective findings in the medical record

(vital signs, standardized nursing assessments, and laboratory results), and thus leaves little room for judgment on the part of the medical record reviewer. We are ultimately unable to determine which direction this potential bias may have influenced our results.

It is also possible that some subjects were treated for presumed sepsis in the ED but the pathway and order set were not utilized, and thus these patients would have been misclassified as negative for physician judgment. The generalizability of our study may be limited, as it was conducted at a large academic children's hospital. Finally, when dissecting algorithm components, it is likely that some components are co-linear (such as heart rate and blood pressure, or blood pressure and capillary refill), thus limiting our ability to fully determine which covariates carry the most weight.

CONCLUSIONS

We provide evidence that an algorithmic approach improves sensitivity of early recognition of severe sepsis/septic shock in an emergency department setting over physician judgment alone, and when used in combination with the more specific physician bedside assessment could improve accuracy of patient identification for appropriate sepsis care.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Lilly CM. The ProCESS trial--a new era of sepsis management. *N Engl J Med*. 2014; 370(18):1750–1. [PubMed: 24635774]
2. Launay E, Gras-Le Guen C, Martinot A, et al. Suboptimal care in the initial management of children who died from severe bacterial infection: a population-based confidential inquiry. *Pediatr Crit Care Med*. 2010; 11(4):469–74. [PubMed: 20068504]
3. de Oliveira CF, de Oliveira DS, Gottschald AF, et al. ACCM/PALS haemodynamic support guidelines for paediatric septic shock: an outcomes comparison with and without monitoring central venous oxygen saturation. *Intensive Care Med*. 2008; 34(6):1065–75. [PubMed: 18369591]
4. Cruz AT, Perry AM, Williams EA, Graf JM, Wuestner ER, Patel B. Implementation of goal-directed therapy for children with suspected sepsis in the emergency department. *Pediatrics*. 2011; 127(3):e758–66. [PubMed: 21339277]
5. Larsen GY, Mecham N, Greenberg R. An emergency department septic shock protocol and care guideline for children initiated at triage. *Pediatrics*. 2011; 127(6):e1585–92. [PubMed: 21576304]
6. Weiss S, Fitzgerald J, Balamuth F, et al. Delayed antimicrobial therapy increases mortality and organ dysfunction duration in pediatric sepsis. *Crit Care Med*. 2014; 42(11):2409–17. [PubMed: 25148597]
7. Thompson G, Macias C. Recognition and Management of Sepsis in Children: Practice Patterns in the Emergency Department. *J Emerg Med*. 2015; 10.1016/j.jemermed.2015.03.012
8. Scott HF, Donoghue AJ, Gaieski DF, Marchese RF, Mistry RD. The utility of early lactate testing in undifferentiated pediatric systemic inflammatory response syndrome. *Acad Emerg Med*. 2012; 19:1276–80. [PubMed: 23167859]
9. Scott H, Deakne S, Woods J, Bajaj L. The prevalence and diagnostic utility of Systemic Inflammatory Response Syndrome vital signs in a pediatric emergency department. *Acad Emerg Med*. 2015; 22:381–9. [PubMed: 25778743]

10. Herasevich V, Kor DJ, Subramanian A, Pickering BW. Connecting the dots: rule-based decision support systems in the modern EMR era. *J Clin Monit Comput.* 2013; 27(4):443–8. [PubMed: 23456293]
11. Nelson JL, Smith BL, Jared JD, Younger JG. Prospective trial of real-time electronic surveillance to expedite early care of severe sepsis. *Ann Emerg Med.* 2011; 57(5):500–4. [PubMed: 21227543]
12. The Algorithm Alert Sepsis Collaborative. [Accessed Sep 1, 2015] Homepage. Available at: <http://quality.pemfellows.com/>
13. Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr Crit Care Med.* 2005; 6(1):2–8. [PubMed: 15636651]
14. Brierley J, Carcillo JA, Choong K, et al. Clinical practice parameters for hemodynamic support of pediatric and neonatal septic shock: 2007 update from the American College of Critical Care Medicine. *Crit Care Med.* 2009; 37(2):666–88. [PubMed: 19325359]
15. Feudtner C, Feinstein JA, Zhong W, Hall M, Dai D. Pediatric complex chronic conditions classification system version 2: updated for ICD-10 and complex medical technology dependence and transplantation. *BMC Pediatr.* 2014; 14:199. [PubMed: 25102958]
16. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988; 44(3):837–45. [PubMed: 3203132]
17. Paul R, Melendez E, Stack A, Capraro A, Monuteaux M, Neuman MI. Improving adherence to PALS septic shock guidelines. *Pediatrics.* 2014; 133(5):e1358–66. [PubMed: 24709937]
18. Embi PJ, Leonard AC. Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study. *J Am Med Inform Assoc.* 2012; 19(e1):e145–8. [PubMed: 22534081]
19. Rhee C, Gohil S, Klompas M. Regulatory mandates for sepsis care--reasons for caution. *N Engl J Med.* 2014; 370(18):1673–6. [PubMed: 24738642]
20. New York State. [Accessed Aug 25, 2015] Hospital Sepsis Protocols. Available at: https://www.health.ny.gov/regulations/recently_adopted/docs/2013-05-01_hospital_sepsis_protocols.pdf
21. Dellinger RP. The Surviving Sepsis Campaign: 2013 and beyond. *Chin Med J (Engl).* 2013; 126(10):1803–5. [PubMed: 23673089]
22. Cruz AT, Williams EA, Graf JM, et al. Test characteristics of an automated age- and temperature-adjusted tachycardia alert in pediatric septic shock. *Pediatr Emerg Care.* 2012; 28(9):889–94. [PubMed: 22929140]

<i>Algorithmic Alert</i>			<i>Physician Judgment</i>			<i>Combination</i>		
	Severe sepsis +	Severe sepsis -		Severe sepsis +	Severe sepsis -		Severe sepsis +	Severe sepsis -
Alert +	81	3,220	PJ +	64	95	Either +	85	3,249
Alert -	7	16,216	PJ -	24	19,341	Both -	3	16,187

Severe sepsis prevalence = 88 (0.45%)

Figure 1.

Two by two tables that demonstrate performance of algorithmic alert, physician judgement, and the combination method.

Table 1

a) Demographics of patients 57 days to 18 years of age with temperature <36 or >38.5 at any point during ED stay or with chief complaint of fever during the 18 month study period. Demographics of the entire ED population during the same period are shown for comparison. B) Characteristics of subjects with positive sepsis screening tests.

a.		
Demographic	Study Cohort	ED Population
Total n	19,524	138,979
Age		
57 days to <1 year	4,411 (22.6)	19,071 (13.7)
1 year to <4 years	9,480 (48.5)	43,073 (31.0)
4 years to <13 years	4,885 (25.0)	52,769 (38.0)
13 years	748 (3.8)	24,066 (17.3)
Sex: female	8,979 (46.0)	65,435 (47.1)
Race		
White	4021 (20.6)	34,291 (24.7)
African American	12,384 (63.4)	87,724 (63.1)
Asian/Indian	937 (4.8)	3,678 (2.6)
Other	2,372 (12.2)	13,318 (9.6)
Disposition: admitted	4,546 (23.2)	27,100 (19.5)

b.				
Patients with Positive Tests	Algorithmic Alert	Physician Judgment	Combination	Sequential
Total n	3,301	159	3,334	126
Age				
57 days to <1 year	203 (6.2)	32 (20.1)	217 (6.5)	18 (14.3)
1 year to <4 years	1,543 (46.7)	35 (22.0)	1,554 (46.6)	24 (19.0)
4 years to <13 years	1,053 (31.9)	65 (40.9)	1,060 (31.8)	58 (46.0)
13 years	502 (15.2)	27 (17.0)	503 (15.1)	26 (20.6)
Sex: female	1,586 (48.1)	63 (39.6)	1,603 (48.1)	46 (36.5)
Disposition				
Admit ICU	327 (9.9)	76 (47.8)	338 (10.1)	79 (62.7)
Admit floor/observation unit	1,527 (46.2)	73 (45.9)	1,571 (47.1)	45 (35.7)
Discharge	1,418 (43.0)	4 (2.5)	1,421 (42.6)	1 (0.8)
Death in ED	1 (0.03)	1 (0.6)	1 (0.03)	1 (0.8)
Other	28 (0.8)	0	19	0

Table 2

Test characteristics of sepsis screening tests.

Test	Algorithmic Alert	Physician Judgment	Combined Method	Sequential Method
	Test Characteristic (95% CI)			
Sensitivity	92.1 (91.67–92.43)	72.73 (72.1–73.35)	96.6 (96.3–96.9)	68.2 (67.5–68.8)
Specificity	83.4 (82.91–83.95)	99.51 (99.41–99.61)	83.3 (82.8–83.8)	99.6 (99.6–99.7)
Positive predictive value	2.5 (2.24–2.67)	40.25 (39.56–40.94)	2.6 (2.3–2.8)	47.6 (46.9–48.3)
Negative predictive value	99.96 (99.93–99.99)	99.88 (99.83–99.93)	99.98 (99.96–100)	99.86 (99.80–99.91)
Positive likelihood ratio	5.6 (5.18–5.95)	148.79 (117.2–1900)	5.8 (5.5–6.1)	200.8 (151.8–266.7)
Negative likelihood ratio	0.09 (0.05–0.19)	0.27 (0.19–0.39)	0.04 (0.01–0.12)	0.32 (0.24–0.43)
Receiver operative characteristic curve area	0.88 (0.85–0.91)	0.86 (0.81–0.91)	0.90 (0.88–0.92)	0.84 (0.79–0.89)

Severe sepsis/septic shock prevalence: 88 (0.45%)

Table 3

A) Description of primary ED diagnoses for the 21 patients identified by the algorithm alert but not identified by physician judgment method who had severe sepsis/septic shock outcome. B) Process and outcome metrics for patients with severe sepsis/septic shock outcome who were vs. were not identified by physician judgment method

A.			
ED Diagnosis	General Inpatient Unit	PICU	Operating Room
Oncology with fever	1	0	0
Pneumonia	2	6	0
Respiratory distress	0	2	0
Pyelonephritis	1	0	0
Abdominal pain	0	1	0
VP shunt infection	0	1	0
Asthma exacerbation	1	0	0
Croup	0	1	0
Altered mental status	0	1	0
Nephrolithiasis	0	1	0
Sepsis/septic shock	0	2	1
Total	5	15	1
B.			
Process/Outcome Metrics	Physician Judgment+ (n=64)	Physician Judgment- (n=21)	p
Time to antibiotics (median minutes, IQR)	27 (1–60)	141 (91–229)	<0.001
PICU LOS (median hours, IQR)	70 (36–108)	228 (93–419)	0.001
Hospital LOS (median hours, IQR)	135 (92–192)	427 (234–650)	<0.001

LOS = length of stay; PICU = pediatric intensive care unit; VP = ventilation/perfusion

Table 4

- a. Algorithmic alert and physician judgment test characteristics across age strata.
- b. Algorithmic alert and physician judgment test characteristics in subjects with at least one comorbidity.

a.						
Method	Age Category	Positive Test n (%)	Severe Sepsis n (%)	Sensitivity	Specificity	Area under ROC curve
Algorithmic alert		Test Characteristic (95% CI)				
	57 days to <1 year (n=4,411)	203 (4.6)	7 (0.2)	85.7 (84.7–86.8)	95.5 (94.9–96.1)	0.95 (0.88–1.0)
	1 year to <4 years (n=9,480)	1,543 (16.3)	19 (0.2)	79.0 (78.1–79.8)	83.6 (83.0–84.6)	0.85 (0.79–0.91)
	4 years to <13 years (n=4,885)	1,043 (21.4)	41 (0.8)	97.6 (97.1–98.0)	79.1 (78.0–80.2)	0.89 (0.87–0.90)
	13 years (n=748)	502 (67.1)	21 (2.8)	95.2 (93.7–96.8)	33.7 (30.3–37.1)	0.65 (0.62–0.68)
Physician judgment	57 days to <1 year (n=4,411)	32 (0.7)	7 (0.2)	85.7 (84.7–86.8)	99.4 (99.2–99.6)	0.92 (0.79–1.0)
	1 year to <4 years (n=9,480)	35 (0.4)	19 (0.2)	63.2 (62.2–64.1)	99.8 (99.7–99.9)	0.81 (0.70–0.93)
	4 years to <13 years (n=4,885)	65 (1.3)	41 (0.8)	75.6 (74.4–76.8)	99.3 (99.1–99.5)	0.87 (0.81–0.94)
	13 years (n=748)	27 (3.6)	21 (2.8)	71.4 (68.2–74.7)	98.4 (97.4–99.3)	0.85 (0.75–0.95)

b.						
Method	Comorbid Condition	Positive Test n (%)	Severe Sepsis n (%)	Sensitivity	Specificity	Area under ROC curve
Algorithmic alert	0 comorbid conditions (n=16,311)	2,030 (12.4)	45 (0.3)	88.9 (88.4–89.4)	87.8 (87.2–88.3)	0.9 (0.87–0.93)
	At least 1 comorbid condition (n=3,213)	1,271 (39.6)	41 (1.3)	95.4 (94.6–96.1)	61.2 (59.5–62.9)	0.79 (0.77–0.81)
Physician judgment	0 comorbid conditions (n=16,311)	91 (2.7)	45 (0.3)	73.3 (72.7–74.0)	99.6 (99.6–99.7)	0.86 (0.80–0.93)
	At least 1 comorbid condition (n=3,213)	68 (2.1)	41 (1.3)	72.1 (70.1–73.6)	98.8 (98.5–99.2)	0.85 (0.79–0.92)

Table 5

Univariate (unadjusted) and multivariable (adjusted) logistic regression to determine association between algorithmic alert components and severe sepsis/septic shock outcome.

Variable	Unadjusted OR (95% CI)	p	Adjusted OR (95% CI)	p
Age	3.2 (2.5–4.2)	<0.001	1.0 (0.8–1.5)	0.7
Female sex	0.7 (0.5–1.1)	0.17	0.6 (0.4–0.9)	0.03
Number of comorbidities	2.3 (1.9–2.8)	<0.001	1.3 (0.9–1.8)	0.14
Hypotension	100.4 (63.3–159.2)	<0.001	13.6 (7.3–25.5)	<0.001
Abnormal heart rate or respiratory rate	10.8 (7.6–15.3)	<0.001	4.8 (3.3–6.9)	<0.001
Physical findings	8.3 (5.6–12.1)	<0.001	5.8 (3.7–9.1)	<0.001
High risk condition	2.9 (2.4–3.6)	<0.001	1.3 (0.9–1.8)	0.22

Physical findings include perfusion or mental status abnormality as recorded in the electronic medical record. High-risk conditions include subjects with a history of malignancy, asplenia, bone marrow transplant, central venous catheter, solid organ transplant, other immunodeficiency, immunocompromised, or immunosuppression. Comorbidity was defined as the presence of at least one complex chronic care condition (Feudtner 2001).¹⁵